

THE DISTRIBUTION OF CYCLE LENGTHS
IN A CLASS OF ABSTRACT SYSTEMS

by

Alan E. Gelfand and Crayton C. Walker

Technical Report No. 236 August 11, 1976

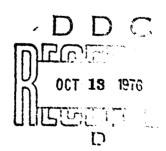
Prepared under Contract NO0014-76-C-0475 (NR-042-267)
Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

Approved for public release; distribution unlimited

DEPARTMENT OF STATISTICS STANFORD UNIVERSITY STANFORD, CALIFORNIA



INTRODUCTION

We suggest that general systems theory is the comparative study of formal models. To us this point seems both obvious and neglected. It is obvious in the sense that general systems theory seeks to comprehend similarities and differences among abstract systems. It is neglected in the sense that study of classes of models has not seen the systematic emphasis that would appear necessary to provide the prospect of a data base complete enough to give theoretical work broad scope.

In this paper we explicitly consider a class of systems defined constructively, that is, not on the basis of the systems' behavior, but in terms of properties of the parts constituting the systems, and the structures which mediate their interactions. After defining the class, we focus attention on a particular behavioral characteristic of the systems, namely the length of the steady-state or cycle length. For the given class of systems we then seek to provide the general systems theorist with a degree of insight into 1) the distributional form of the cycle length considered as a random variable under random structural and initial-state variation, and 2) the effect of system size directly on (average) cycle length and on other parameters of these distributional models.

THE SYSTEMS EXAMINED

The particular class of systems we consider is the set of all systems which are individually fixed in structure and built up from functionally identical elements having two inputs and two internal states. Additionally we assume our systems are not influenced by the environment. The elements' next internal state is a determinate

function of both the internal states and the input values existing at each element at the present instant of idealized system time. The elements' outputs carry the various elements' internal states. In summary, our systems describe a subset of binary, autonomous, fixed structure, deterministic, automata.

We have already outlined our reason for choosing a class of systems for study—our conviction in the appropriateness of this approach for the development of general systems theory. We will now try to justify the particular choice made. In that connection it is interesting that while our biases toward study of relatively narrow classes of systems stem from our perception of the field of general systems theory, these biases are consistent with the research emphases of some late automata theoretic writing [4,5].

In studying the defined class of systems we are in effect recommending this class to the general systems theorist as a worthy object of attention. There are, as we have mentioned, reasons related to research strategy for this, but there exist substantive reasons as well. Among fixed-structure models built of functionally heterogeneous elements with the same number of inputs per element (K), Kauffman [3] argues that biological applicability is maximized for K = 2. The criteria used by him for biological relevance, namely, relatively localized steady-state behavior, resistance of steady-state behavior to disturbance, and restricted possibilities of change among the different steady-states available have wider significance. The first criterion can be considered a demand that a system's steady-state at the very least be empirically recognizable. The second criterion requires the

system to be behaviorally robust. The third criterion requires the system to have constrained developmental opportunities. The second and third criteria are clearly desirable in models of many biological systems, and can be seen to apply in other areas as well. The important first criterion appears close to a necessary condition in models for any real-world system to which the term steady-state behavior has empirical significance. Of course there are qualifications; it is not asserted that a model of any real world phenomenon must be a network of heterogeneous elements with two inputs. Rather we take Kauffman's analysis as demonstrating the theoretically pivotal nature of the class of network models for which K = 2.

Our own class of model systems differs from Kauffman's in having functionally homogeneous elements (all elements in any one system are identical) with two internal states rather than one. This class can be finally justified by pointing to the value for theoretical development of determining the behavior of simple, functionally "pure" systems, leaving the study of the behavior of mixtures of elements (as in Kauffman's systems) as a natural next step. We should also point out that elements with one internal state are included in our class.

The functional description of an element can be expressed in general in the following form given in Figure 1, where entries a,b,...,h are the next internal state, and next output state of the element. They take on values 0 or 1 in any specific table. We will refer to a specific table by interpreting the entries abcd and efgh as two binary integers in the notation T(abcd,efgh). A system using the table T(1001,0110) would be referred to as a T(9,6) system. The defining table will also be called a transformation.

FIGURE 1: Matrix Form of a Transformation

Present Internal State

			_		0	1
			0	0	a b c	c
.	- .	.	0	1	ъ	f
Present	Input	States	1	0	c	g
			1	1	d	h

There are $2^8 = 256$ different transformations. Therefore, there are 256 functional types of systems in our class. Where attention is restricted to length of behavior (steady-state or transient) these 256 types reduce to the 88 behavioral equivalence classes [6,7] for which data are taken in this paper.

A specific system is formed by taking N elements which realize the same transformation T and joining them so that all elements' inputs are connected to outputs. While the behavior of that specific system is observed, no changes are made in the structure (the pattern of connections) or its function (T). We follow convention in referring to the set of internal states of the elements of a given system as the system state. It is clear that there are exactly 2^N distinct system states. It is also clear that if the system is started at an arbitrary system state, it will step through a (possibly void) set of transient states before reaching its steady-state, a set of states through which it will continuously cycle. The number of states in this cycle is the cycle length V, $1 \le V \le 2^N$. By system size we mean the value of N.

PROCEDURE

As proposed in the introduction, we will now develop a model for the distribution of cycle lengths. In the process we will garner considerable insight into how system size affects cycle length. First of all, a clarification with respect to the notion of a distribution of cycle lengths is in order. For a given system size N , a given transformation T, a fixed system starting state R, and a fixed set of system connections C, the cycle length V is then specified in a completely deterministic manner, as described in the previous section. i.e., V = G(R,C,N,T). However at a fixed N, the number of possibilities for choosing R is 2 and the number of possibilities for choosing C is N^{2N} so that the process of investigating the function G for each of 88 choices of T with varying choices of N is computationally infeasible. Hence to make this examination tractable we resort to an artifice; R and C are selected at random, i.e., such that each possible starting vector R is equally likely and has a 1/2N chance of selection and each choice of connections C is equally likely with 1/N^{2N} chance of selection. Since R and C are random, V will now be random as well and a distribution of V exists in that sense. For N = 2 the exact distribution of V can be readily obtained given T. This case will be considered shortly. The problem of calculating the exact distribution of V for N = 3 or 4 is immense, and for larger N is intractable. Moreover, we do not intend to attempt a transformation of variables to obtain V's distribution since the function G at a fixed N and T appears hopeless to describe. Nor is it clear for our purposes that we necessarily want the exact distribution

of V at each N and T. Rather, since V's distribution now depends only on the parameters N and T, we would prefer to describe a family of density functions $f_V(v;N,T)$ which seem to effectively describe V's behavior and which over varying T's allow us to examine the dependence of V's distribution on N.

Before we delve more deeply into this question we pause to briefly consider the case for N=2. In this case there are 4 possible starting configurations and 16 possible connection sets, so that there are but 64 possible R and C choices, each with likelihood 1/64 of being selected. The sample space for V is $\{1,2,3,4\}$. As an example, the transformation T(9,6) was examined and yielded P(V=1)=7/8, P(V=2)=1/8, P(V=3)=P(V=4)=0.

The remaining 87 distinct transformations could similarly be examined, but there seems to be little gain in this. Instead we turn to the more interesting situation with general N. Employing the notions of randomness described above and using the computer, we obtained 100 observations from each of the cycle length distributions given by the 88 distinct transformations taken with system sizes $N = 4,5,\ldots,17$, i.e., we examined 100 observations from each of $88 \times 14 = 1232$ distributions. While 100 observations is hardly enough to develop a highly accurate picture of each of these distributions, it is enough to reveal their general behavior, particularly with regard to varying N, which, as we mentioned above, is a primary concern. A preliminary examination of our results indicates that the 88 transformations can be more or less classified into two major groups; those which are essentially degenerate over N and those which exhibit changing, often rather volatile, behavior

over N. The former group is largely made up of transformations which exhibit almost trivial behavior in the sense of yielding cycle lengths which are virtually always one, virtually always two, or a relatively stable proportion of one's and two's regardless of N . The latter group contained transformations which clearly seemed to be varying (increasing) in N and as N grew larger occasionally resulted in cycle lengths of the order of 104. We note that at a given N the maximum cycle length* 2^{N} and is of order 10^{4} for N > 14. Moreover, for these more volatile transformations several cycles lengths on the order of the maximum were observed even though our samples were only size 100, indicating that these lengths are not rarities but rather occurrences which may be expected with some degree of regularity. Additionally for these more interesting transformations, the maximums of the samples seemed to be increasing exponentially in N analogously to the maximum cycle length. These observations, coupled with an effort to achieve a more manageable range for the data, led to a decision to make a logarithmic transformation on the observations. The base of the log transformation is, of course, not important since it provides nothing more than a scaling of the observations. For convenience, base ten was employed. Henceforth in all discussion of cycle length we will implicitly refer to log cycle length. Although the sample spaces of the observations are clearly discrete, as a result of the log transformation and in the interest of mathematical convenience we attempted to fit a continuous family of distributions to the data. Noting that the distributions are bounded, and as they appeared

[&]quot;The maximum cycle length actually depends on T as well and can never exceed 2^N-2 , $N\geq 2$, for any T .

to be more or less unimodal, we were naturally led to the Pearson Type II or generalized Beta family of distributions. Although these distributions are ordinarily parametrized by four parameters, as a result of our data transformation it is clear that the lower limit of the sample space ought to be taken as zero. Thus we examined the family of three parameter distributions whose probability density function is given by

$$f_{V}(v) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{v^{\alpha-1}(\gamma-v)^{\beta-1}}{\gamma^{\alpha+\beta}}$$
 $0 < v < \gamma$

with $\alpha > 0$, $\beta > 0$, $\gamma > 0$.

The first three moments of this distribution are

$$\mu_{1} = \frac{\alpha}{\alpha + \beta} \gamma ,$$

$$\mu_{2} = \frac{\alpha(\alpha + 1)\gamma^{2}}{(\alpha + \beta)(\alpha + \beta + 1)} ,$$

$$\mu_{3} = \frac{\alpha(\alpha + 1)(\alpha + 2)\gamma^{3}}{(\alpha + \beta)(\alpha + \beta + 1)(\alpha + \beta + 2)} ,$$

and

$$E(V) = \mu_1$$
, $var(V) = \frac{\alpha\beta\gamma^2}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

If m_1, m_2, m_3 are the first three sample moments, then we can obtain three equations in three unknowns to solve for sample estimators of α , β , γ . The equations are gotten by replacing μ_i by m_i , i=1,2,3 above. The solutions are

$$\hat{\alpha} = \frac{m_1 (m_1 \hat{\gamma} - m_2)}{(m_2 - m_1^2) \hat{\gamma}},$$

$$\hat{\beta} = \frac{(m_1 \hat{\gamma} - m_2) (\hat{\gamma} - m_1)}{(m_2 - m_1^2) \hat{\gamma}},$$

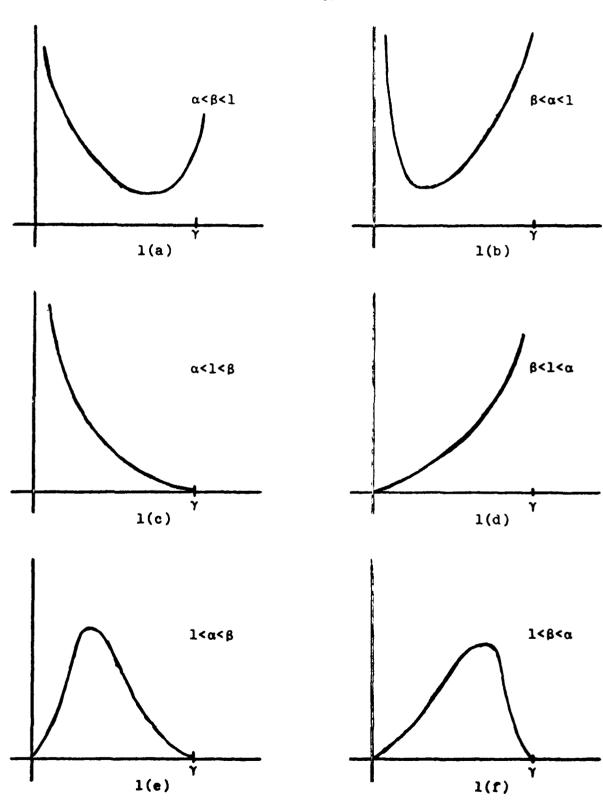
where

$$\hat{\gamma} = \frac{m_1 (m_2^2 - m_3 m_1) + m_3 (m_2 - m_1^2)}{(m_2^2 - m_3 m_1) + m_2 (m_2 - m_1^2)}.$$

Varying α , β and γ varies the skew and range of these distributions. The pictures in Figure 2 describe the sorts of behavior possible for the density function, $f_{\gamma}(v)$. Increasing γ , of course, primarily expands the measurement scale although all the moments of the distribution are increased. Changing α and β varies the shape of distribution. For a fixed γ increasing α (β) relative to β (α) increases (decreases) the mean. By examining the variance expression it is apparent that if either α or β or both grow large the variance or spread will be decreased; but also if either α or β or both are made very small the variance (spread) will be decreased. Thus for a fixed γ making α (β) large relative to β (α) increases the area to the right (left) of the scale and making α (β) small relative to β (α) increases the area to the left (right) of the scale. As a result a distribution for which a is very small relative to β will be essentially degenerate at 0 (an extreme case of 1(a), 1(c) or 1(e) in Figure 2). Similarly a distribution for which α is very large relative to β will be essentially degenerate at γ (an extreme case of 1(b), 1(d) or 1(f) in Figure 2). A distribution such that α and β are quite large but of the same

FIGURE 2: Pearaon Type II Curves

ard



order of magnitude will be essentially degenerate at $\alpha\gamma/\alpha+\beta$. Further discussion of the properties of these curves is available in the reference texts [1,2].

RESULTS

Let us turn to an examination of our cycle length distributions in terms of the above family of distributions. Those distributions which tend to be active (i.e., nondegenerate) were usually fit by a curve of the form 1(e) or 1(e) (occasionally 1(f)) from Figure 1. The essentially degenerate distributions were usually fit by extreme cases of 1(a) or 1(c) (where the cycle length was almost always 1 so that the log cycle length was almost always 0) or by extreme cases of 1(e) or 1(f), i.e., with α , β large so that the variance is small (where, for example, the cycle length was almost always 2 so that the log cycle length was almost always .3010). There are several distributions which do not seem to be at all well described by the model. The poor fit may be explained by one or more of the following reasons.

- (i) We are fitting a family of continuous distributions to observations which are clearly from a discrete scale.
- (ii) We have only n = 100 observations for each empirical distribution--hardly enough to insure that our empirical picture is close to the true underlying distributional picture.
- (iii) For some transformations, our proposed distributional family may be inappropriate. For these transformations the true distribution may be too "discrete" (i.e., the

possible log cycle lengths may not be dense enough within the range) or may not be unimodal (i.e., the generalized Beta family defines only unimodal distributions).

Nonetheless for most of the transformations the results were rather gratifying and indicate at least to some extent the appropriateness of our model. In fact 70 of the 88 transformations taken over N seemed to be well fit. Only 18 were badly described and these 18 might be fit better with larger sample sizes. Table 1 considers the behavior of several of the transformations at various N's to illustrate the above discussion.

We now return to the more pertinent question regarding the relationship between system size and cycle length (equivalently log cycle length). We first note that the question of goodness of fit of the distributional models is separate from that of the behavior of cycle length with increasing sample size. To further clarify this point there are four possible situations as in Figure 3.

FIGURE 3: A Classification of Transformations

	Cycle Length Behavior is Strongly Related to N	Cycle Length Behavior is Unrelated to N
Model Fits Transformation Over N	I	11
Model Does Not Fit Transforma- tion Over N	III	IV

TABLE 1 Describing the distributions of various transformations

Transformation	Description	Z	4 8	(02	\ \ \	Curve Type (From Figure 2)
1(0,0)	Degenerate at log l=0	any	0	0	0	limiting case of 1(a)
T(15,0)	Degenerate at log 2=.3010	any	4.65×104	3.08×10 ⁴	ιņ	limiting case of i(e) or l(f)
T(6,9)	Active Large Range	10	1.21	1.36	2.17	1(e)
T(6,9)	Active Large Range	15	1.92	ተተ ፣	3.38	1(£)
T(6,1)	Active Lesser Range	10	0.86	2.42	1.54	1(c)
T(6,1)	Active Lesser Range	16	1.18	2.55	1.76	1(e)
T(1,7)	Essentially Degenerate at log l=0	œ	1.01×10 ⁻⁸	7.71×10 ⁻⁷	.3010	1(a)
T(1,7)	Essentially Degenerate at log 1=0	15	1.01×10 ⁻⁸	7.71×10 ⁻⁷	.3010	1(a)
T(10,2)	Model does not fit well; bi-modal at log 1 and log 2	Φ	-23.35	1.85	.08	1
T(10,2)	Model does not fit well; bi- modal at log l and log 2	15	4.09	-15.46	-1.26	

We expect that for most transformations increasing N will in some sense increase cycle length. But it is also apparent that there will exist transformations which will produce cycle lengths which will be independent of N. That is, certain transformations by their definition will result in short cycle lengths regardless of system size.

As a result there will be transformations in each of the four categories. Of course the column classification is not perfectly dichotomous. While most of the 88 transformations could be comfortably classified in one of the two categories, there were perhaps nine transformations which required more data to conclude exactly where to place them. Not surprisingly, most of these arose in situations where the model did not fit the transformation well. Table 2 summarizes the categorization of the transformations, including plausible placements for the above nine.

If we wish to focus directly on the dependence of cycle length on system size, it would make sense to examine and correlate with system size certain characteristics of our samples such as the sample mean, sample median, sample standard deviation and sample maximum. Although the data are not continuous, a product moment correlation is likely to be as effective as any other measure of association in this case. A large positive correlation value will support the contention of strong dependence of cycle length on system size while a correlation value near zero (positive or negative) will indicate relative independence of cycle length and system size. We would not expect large negative correlations.

More specifically, if for a particular transformation cycle length is truly dependent upon system size then, in terms of the artificial randomness we have created, the distribution mean ought to be strongly

TABLE 2
A categorization of the transformations

I: (1,6), (1,8), (1,9), (1,10), (1,11), (1,14), (2,0), (2,13), (2,14), (3,0) (2,4), (2,5), (2,8), (2,8), (2,9), (2,10), (2,12), (2,13), (2,14), (3,0) (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (3,8), (3,9), (3,10), (3,12) (3,14), (6,0), (6,1), (6,2), (6,6), (6,8), (6,9), (6,10), (6,14), (7,2) (7,6), (7,10), (7,14), (8,6), (9,6), (9,10) (9,10), (0,11) (0,11), (0,12), (1,0), (1,1), (1,2), (1,3), (1,7), (2,11), (8,10), (8,14) (8,14), (14,9), (14,8), (15,0)	(7,1), ((10,12), ((7,0), (7,0), (
--	-----------------------------------	--	--

correlated with N. This should be true of the distribution median as well, although perhaps not quite so strongly since for discrete distributions it may turn out that the median is monotonically increasing in N but not strictly. Further, if cycle lengths are affected by N we might reasonably expect the upper bound on the sample space of the distribution (which, of course, is not necessarily the maximum possible cycle length at a given N) to be strongly correlated with N. Finally, the distribution variance need not necessarily be strongly correlated with N . Even though increasing N leads to observing increasingly larger cycle length measurements, the distributional spread as characterized by variance may or may not be increasing. Additionally, since the variation is artificially induced (by our arbitrary choice of an equiprobable sampling regime) it is questionable as to how much interpretation may be attached to it. Hence the pertinence of this final correlation measurement to the question of the dependence of cycle length on N is not assured. However, if the population variance is strongly correlated with N, of necessity the space of observations must be increasing in N so there will be some evidence of dependence. As a result, the sample estimates of all of these distributional characteristics ought to reflect these expectations as well.

Suppose in addition we are in Situation I and wish to examine the effect of increased system size on the parameters of distributional model. As we have discussed, the upper bound on the log cycle length space ought to be increasing and in fact the mass of the distribution ought to be moving to the right as well. This should be reflected in a parametric dependence on N . In light of discussion in the previous section, an

increase in Y will achieve both these effects but additionally an increase in α will also adjust the skew of the distribution. We would not expect β to increase much if at all. As a result we would expect $\hat{\gamma}$ to be most strongly correlated with N , α next and then β . This was almost universally borne out by the data. In Situations II and IV we would expect little correlation of N with any of our sample characteristics or parameter estimates. Again this was usually true, although for an occasional transformation a surprisingly large correlation value with $\hat{\alpha}$ or $\hat{\beta}$ was obtained. After more careful examination it became apparent that these values were obtained upon correlating N with estimates on the order of 10⁻⁶ or smaller. These estimates, although perhaps slowly increasing in $\, \, {\rm N} \,$, are still so small that the cycle length distribution remains essentially degenerate over N and hence is still essentially independent of N. In Situation III we expect the parameter estimates to be weakly correlated with N while the sample characteristics ought to be strongly correlated with N. Table 3 examines typical transformations in each of the above situations and supports the preceding discussion.

Hence it is possible to conclude that for certain transformations there is highly convincing evidence that cycle length is directly related to system size. For others there is convincing evidence that essentially no relationship exists and for very few is the evidence inconclusive. Roughly 2/3 of the transformations are in the first case, approximately 1/4 in the second and the remaining 1/12 in the last case. Again perhaps additional data would allow us to resolve the ambiguities in the last case.

Correlating cycle length and system size for various transformations TABLE 3

				200	Correletton with	4 5		
Transformation	Situation	۲,	⟨ ၓ	, 68 68	* ***	* *	*	**
T(6,9)	н	.973	746.	.772	966.	.986	.992	066.
T(7,6)	н	.910	.801	.563	.992	956.	.843	026.
T(0,0)	II	0	0	0	0	0	0	0
T(1,7)	II	108	244.	447	217	0	141	114
T(10,4)	III	197	991.	.038	986.	.926	.980	. 950
T(11,4)	III	304	146	344	.993	.975	.962	096.
T(7,0)	ΙΛ	526	098	001	.633	.002	ħ06·	.232
T(11,2)	IV	.275	199.	378	445	.002	158	127

sample mean sample median sample standard deviation sample maximum ***

....

REFERENCES

- [1] W. P. Elderton and N. L. Johnson, <u>Frequency Curves</u>, Cambridge University Press, London, 1969.
- [2] N. L. Johnson and S. Kotz, <u>Continuous Univariate Distributions</u>, Houghton-Mifflin, Boston, 1971.
- [3] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," <u>Journal of Theoretical Biology</u>, 22, 1969, pp. 437-467.
- [4] R. McNaughton and S. Papert, <u>Counter-Free Automata</u>, M.I.T. Press, Cambridge, Massachusetts, 1971.
- [5] M. Minsky and S. Papert, <u>Perceptrons</u>, M.I.T. Press, Cambridge, Massachusetts, 1969.
- [6] C. C. Walker, "Behavior of a class of complex systems: the effect of system size on properties of terminal cycles," <u>Journal of Cybernetics</u>, 1, 1971, pp. 55-67.
- [7] C. C. Walker and A. A. Aaydryan, "Amount of computation preceding externally detectable steady-state behavior in a class of complex systems," Journal of Bio-Medical Computing, 2, 1971, pp. 85-94.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)	
REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
REPORT NUMBER 2 GOVT ACCESSION NO.	
236	
4 TITLE (and Sublisio)	TYPE OF BERGET & BERGO COVERED
THE DISTRIBUTION OF CYCLE LENGTHS IN A	Technical Report,
CLASS OF ABSTRACT SYSTEMS.	6 PERFORMING ORG. REPORT NUMBER
7. AUTHOR(e)	
AUTHORIS)	CONTRACT OR GRANT NUMBER(s)
Alan E./Gelfand Crayton C./Walker	NAAAA - 76-C-8475
9 PERFORMING ORGANIZATION NAME AND ADDRESS	16 PROGRAM ELEMENT PROJECT, TASK
Department of Statistics Stanford University	NR-#42-267
Stanford, Calif. 94305	(1117,042-207)
11 CONTROLLING OFFICE NAME AND ADDRESS	12. RERORT DATE
Office of Naval Research Statistics & Probability Program Code 436	11 Augu 176
Arlington, Virginia 22217	21
18 MONITORING AGENCY NAME & ADDRESS(II different from Controlling Office)	18. SECURITY CLASS. (of this report)
1	UNCLASSIFIED
	The DECLASSIFICATION/DOWNGRADING
	SCHEDULE
16 DISTRIBUTION STATEMENT (of this Report)	(12)
	(2) 000
Approved for Public Release; Distribution Unlimi	ted dap.
	(
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 29, If different to	Record
	4TR-236\
18. SUPPLEMENTARY NOTES	
	[
	ļ
19. KRY WORDS (Continue on reverse side if necessary and identify by block number)	
Cycle length, system size, generalized Beta dist	aibutions simulation
office rength, system size, generalized beta disti	indrions, Simulation
	1
28. ABSTRACT (Continue on reverse side if necessary and identify by block number)	
Systems which are built up from functionally idea	itical alemante haudne tun
inputs and two internal states are subjected to i	andom initial positions and
random structural connections such that a random	determination of cycle
length is achieved. This artifice enables tracts	able examination of the
systems over varying system size. In particular distribution of the cycle length variable and al	we present a model for the
system size on cycle length variable and all	the errect of
DD 1 JAN 72 1473 EDITION OF 1 NOV 66 18 OF CUTE	UNCLASSIFIED # 13 presents
MEUNITY CLA	MIFICATION OF THIS PAGE (When Bate Sheered)
is considered,	332580
	338300

人B